Diya Das, PhD

diyadas@alumni.princeton.edu | linkedin.com/in/diyadas | web: diyadas.github.io

Skills and Knowledge

- generation of biological models from genomic data
- single-cell RNA-sequencing analysis
- assay for transposase-accessible chromatin sequencing (ATAC-seq) and analysis
- assorted molecular biology lab techniques (FACS, PCR, in situ hybridization, etc.)

Tools

- High performance computing
- R and RStudio
- Python (*pandas*, *dask*)
- MATLAB
- Bash
- Git and GitHub

Experience

Development Sciences Informatics, Genentech; Informatics Analyst II — June 2019 – Present

Lab of John Ngai, UC Berkeley

Postdoctoral Researcher — June 2018 – May 2019

Graduate Student Researcher — May 2013 – May 2018

- Developed computational pipelines for analyzing Fluidigm and 10X Genomics single-cell RNA-sequencing data (R/Bioconductor packages *scone, zinbwave*, *Seurat, clusterExperiment* and *slingshot*).
- Identified when cells choose between neuronal and non-neuronal fates by applying resampling-based ensemble clustering and lineage trajectory algorithms developed in collaboration with statisticians and computer scientists (packages *clusterExperiment* and *slingshot*).
- Explored gene regulation by chromatin accessibility in stem cells using ATAC-seq (Python; *dask*, *deepTools*).
- Wrote scripts to retrieve sequencing data and accompanying experimental metadata from NCBI databases (available at github.com/diyadas/utils; uses Python package *xml* and NCBI's *sra-toolkit*)
- Code for two co-first author papers at github.com/diyadas/HBC-regen & github.com/rufletch/p63-HBC-diff. Shiny app available at github.com/diyadas/olfactory-expression.

CDIPS Data Science Workshop, UC Berkeley; Participant — July 2016

- Developed topic tree for relationships between Wikipedia articles (Python; packages *BeautifulSoup*, *lxml*).
- Code available at github.com/diyadas/Topic-Ontology.

Department of Molecular & Cell Biology, UC Berkeley; Graduate Student Instructor — 2013 – 2015

- Coordinated curriculum for discussion sections and review sessions, exam grading, and course policies with three professors and graduate student instructors (GSIs) for two undergraduate courses.

Lab of Sam Wang, Princeton University; Undergraduate Researcher — June 2009 – June 2012

- Analyzed canine temperament and neuroanatomy with MRI, culminating in award-winning thesis (MATLAB).

Education

University of California, Berkeley — August 2012 – May 2018

PhD in Molecular & Cell Biology

Princeton University — September 2008 – June 2012

AB in Molecular Biology

Certificate in Neuroscience, Quantitative and Computational Neuroscience honors track

Fellowships

- Moore/Sloan Data Science Fellow, Berkeley Institute for Data Science — August 2017 – May 2019
- Elizabeth Roboz Einstein Fellow in Neurosciences & Human Development — January 2015 – May 2015
- California Institute for Regenerative Medicine Predoctoral Fellow — January 2015 – December 2015

Diya Das diyadas@alumni.princeton.edu

Publications (* = co-first author publication) full list: diyadas.github.io/publications

- D. Sholler*, **D. Das***, F. Hoces de la Guardia* et al. (2019). Best Practices for Managing Turnover in Data Science Groups, Teams, and Labs. SocArXiv: https://doi.org/10.31235/osf.io/wsxru.
- D. Risso, L. Purvis, R. Fletcher, **D. Das**, J. Ngai, S. Dudoit and E. Purdom. (2018). clusterExperiment and RSEC: A Bioconductor package and framework for clustering of single-cell and other large gene expression datasets. PLOS Computational Biology *14*, e1006378.
- R.B. Fletcher, **D. Das** and J. Ngai. (2018). Creating Lineage Trajectory Maps Via Integration of Single-Cell RNA-Sequencing and Lineage Tracing. BioEssays *40*, 1800056.
- K. Street, D. Risso, R.B. Fletcher, **D. Das**, J. Ngai, N. Yosef, E. Purdom and S. Dudoit. (2018). Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics *19*, 477.
- L. Gadye*, **D. Das***, M.A. Sanchez*, K. Street, A. Baudhuin, A. Wagner, M.B. Cole, Y.G. Choi, N. Yosef, E. Purdom, S. Dudoit, D. Risso, J. Ngai and R.B. Fletcher. (2017). Injury Activates Transient Olfactory Stem Cell States with Diverse Lineage Capacities. Cell Stem Cell *21*, 775-790.e9.
- R.B. Fletcher*, **D. Das***, L. Gadye, K.N. Street, A. Baudhuin, A. Wagner, M.B. Cole, Q. Flores, Y.G. Choi, N. Yosef, E. Purdom, S. Dudoit, D. Risso and J. Ngai. (2017). Deconstructing Olfactory Stem Cell Trajectories at Single-Cell Resolution. Cell Stem Cell *20*, 817-830.e8.

Talks and Tutorials full list: diyadas.github.io/presentations

- **D. Das**. Unraveling Adult Tissue Regeneration. 2018 Moore-Sloan Data Science Environments Summit, Park City, Utah. October 10, 2018.
- **D. Das**. Unraveling Tissue Regeneration with Single-Cell RNA-Sequencing. Northern California Computational Biology Symposium, UCSF. October 6, 2018.
- **D. Das**, K. Street and D. Risso. Analysis of single-cell RNA-seq data: Dimensionality reduction, clustering, and lineage inference. BioC 2018, Toronto, Ontario. July 27, 2018.
- **D. Das**. Injury Activates Transient Olfactory Stem Cell States with Diverse Lineage Capacities. UC Berkeley Developmental & Regenerative Biology Retreat. November 14, 2017.
- **D. Das**. Deconstructing Olfactory Stem Cell Trajectories at Single-Cell Resolution. UC Berkeley Developmental & Regenerative Biology Retreat. January 9, 2017.
- C. Cypranowska and **D. Das**. Intro to Genomics Data Wrangling (Data Carpentry Workshop). Aug 6-7, 2018.
- **D. Das**, R. Barter and R. Barnes. Intro to Shell, Git and R (Software Carpentry Workshop). June 11-12, 2018.
- **D. Das**, et al. Various tutorials on Bash, GitHub, and R. UC Berkeley. 2017-2018. Code available online: github.com/diyadas/bash-tutorial, github.com/diyadas/yagt, github.com/diyadas/tutorials.

Leadership full list: diyadas.github.io/leadership

Berkeley Institute for Data Science:

Best Practices and Meta-Research Working Group October 2018 – May 2019
Executive Committee, Fellow Representative June 2018 – November 2018

Beyond Academia:

Co-Director and Development Lead January 2017 – January 2018

- Organized recruiting and personally held 1-1 informational meetings with 10 prospective members; 6 joined.
- Managed partnerships with campus units and planned on-campus recruiting event for employers.
- Planned two 2-day annual conferences for 300+ peers on career options outside academia, featuring 100+ speakers, with team of ~20 graduate students and postdocs (started May 2016).

Logistics, Speakers and Development Committees May 2016 – March 2017

- Developed conference schedule for 32 workshops/panels based on 100+ speaker availabilities.
- Organized and coordinated speakers for four panels.
- Redefined advisory board responsibilities and membership, selecting new advisors to fulfill needed expertise.

CDIPS Data Science Workshop, **Co-Director** January 2017 – August 2017
MCB Graduate Student & Alumni Association, **Student Co-President** June 2015 – May 2017
MCB Graduate Affairs Committee, **Student Representative** August 2013 – May 2015
Student Health Advisory Committee, **Grad Student Representative** September 2012 – April 2016
Expanding Your Horizons at Berkeley, **Finance Agent and Signatory** September 2012 – April 2015